Appendix

Description of Fit Statistics Shown in Table 1

The average posterior probability for group j is $AvePP_j$. If individuals are assigned to distinct groups with no ambiguity, the $AvePP_j$ would be 1.0 for each group—the closer the $AvePP_j$ are to 1.0, the better the model fit. An AvePP greater than 0.7 for all groups is recommended (Nagin, 2005). In this study, the observed $AvePP_j$ are much greater than 0.7 (and are actually approaching 1.0), suggesting that subjects with different heroin use trajectories can be very accurately placed into a trajectory group. The lowest AvePP across the five groups is 0.990173.

The odds of correct classification for a trajectory group (OCC_j) is $(AvePP_j/(1- AvePP_j))/(\pi_j/(1-\pi_j))$, where π_j is the is the population size of trajectory group j. The numerator is based on the maximum probability rule while the denominator is based on random assignment. The OCC would equal 1.0 for a given trajectory group if the maximum probability rule is not better than random guessing. Nagin (2005) recommends an OCC of greater than five for each group. In this study, OCC_j is much greater than five, with the lowest at 310.04.

The final model fit statistic is the difference between estimated group probabilities π_j and the proportion P_j assigned to the group using the maximum probability rule. Good model fit is indicated when these two quantities are similar for each group j (Nagin, 2005). The probability of group membership (as estimated from the model) and the proportion assigned to each group using the maximum probability rule are almost identical for each group in this study.

Model Fit Statistics for Two-Group through Six-Group Models with Higher Order Polynomials

The model fit statistics for the estimated models with two, three, four, five, and six discrete groups are presented in Table A1. Higher order polynomials (cubic) were used to determine the appropriate number of groups. The AvePP, OCC, and estimated proportion for each group j were estimated using Stata programming code written by Andrew Wheeler, University of Texas at Dallas (https://andrewpwheeler.wordpress.com/2016/10/06/group-based-trajectory-models-in-stata-some-graphs-and-fit-statistics/). The code is reproduced below.

```
program summary_table_procTraj
    preserve
    *average posterior probability
    gen Mp = 0
    foreach i of varlist _traj_ProbG* {
        replace Mp = `i' if `i' > Mp
    }
    sort _traj_Group
    *odds of correct classification
    by _traj_Group: gen countG = _N
    by _traj_Group: egen groupAPP = mean(Mp)
    by _traj_Group: gen counter = _n
    gen n = groupAPP/(1 - groupAPP)
    gen p = countG/ _N
    gen d = p/(1-p)
    gen occ = n/d
    *estimated proportion for each group
```

```
    scalar c = 0
    gen TotProb = 0
    foreach i of varlist _traj_ProbG* {
       scalar c = c + 1
       quietly summarize `i'
       replace TotProb = r(sum)/ _N if _traj_Group == c
    }
    list _traj_Group countG groupAPP occ p TotProb if counter == 1
    restore
end
```

The BIC_1 and BIC_2 indicate increasing model fit from the two-group model to the six-group model. In each case, the additional model fit statistics are in the appropriate range. We decided that the five-group model was more appropriate than the six-group model because the graphs (Figures A2 and A3) show that the five-group model fits the data better.

Table A1. Model Fit Statistics with Higher-Order (Cubic) Polynomials

Model	n	BIC_1	BIC_2	$AvePP_j$	OCC_j	P_j	π_j	difference between P_j and π_j
Two Groups		-9455.84	-9442.30					
Group 1	103			0.999980	55309.00	0.485849	0.486037	-0.000189
Group 2	109			0.999615	2453.57	0.514150	0.513962	0.000189
Three Groups		-8716.55	-8694.88					
Group 1	62			0.999898	23748.33	0.292453	0.292462	-0.000009
Group 2	59			0.999224	3337.93	0.278302	0.278294	0.000007
Group 3	91			0.999584	3191.938	0.429245	0.429244	0.000002
Four Groups		-7981.75	-7951.96					
Group 1	67			0.999972	77086.80	0.316038	0.316052	-0.000015
Group 2	35			0.997480	2002.01	0.165094	0.164865	0.000229
Group 3	43			0.999466	7359.37	0.202830	0.203255	-0.000425
Group 4	67			0.997775	970.35	0.316038	0.315828	0.000210
Five Groups		-7519.00	-7481.09					
Group 1	28			0.998904	5986.27	0.132076	0.132522	-0.000446
Group 2	66			0.995733	516.19	0.311321	0.309997	0.001324
Group 3	25			0.999964	206057.50	0.117925	0.118682	-0.000757
Group 4	56			0.998833	2384.29	0.264151	0.264391	-0.000240
Group 5	37			0.996913	1527.22	0.174528	0.174409	0.000119
Six Groups		-7167.45	-7121.42					
Group 1	23			0.999788	38707.02	0.108491	0.108479	0.000012
Group 2	61			0.999961	63017.98	0.287736	0.288421	-0.000685
Group 3	28			0.996674	1968.92	0.132076	0.131895	0.000181
Group 4	40			0.994393	762.63	0.188679	0.187884	0.000795
Group 5	26			0.994485	1289.97	0.122642	0.121965	0.000676
Group 6	34			0.998806	4379.65	0.160377	0.161356	-0.000979

n = sample size for the group; BIC_1 = 3,180 observations; BIC_2 = 212 individuals; $AvePP_j$ = average posterior probability for group j; OCC_j = odds of correct classification; P_j = proportion assigned using the maximum probability rule; π_j = estimated group probabilities

Figure A1. Model Estimated with Mean Yearly Frequency of Heroin Use

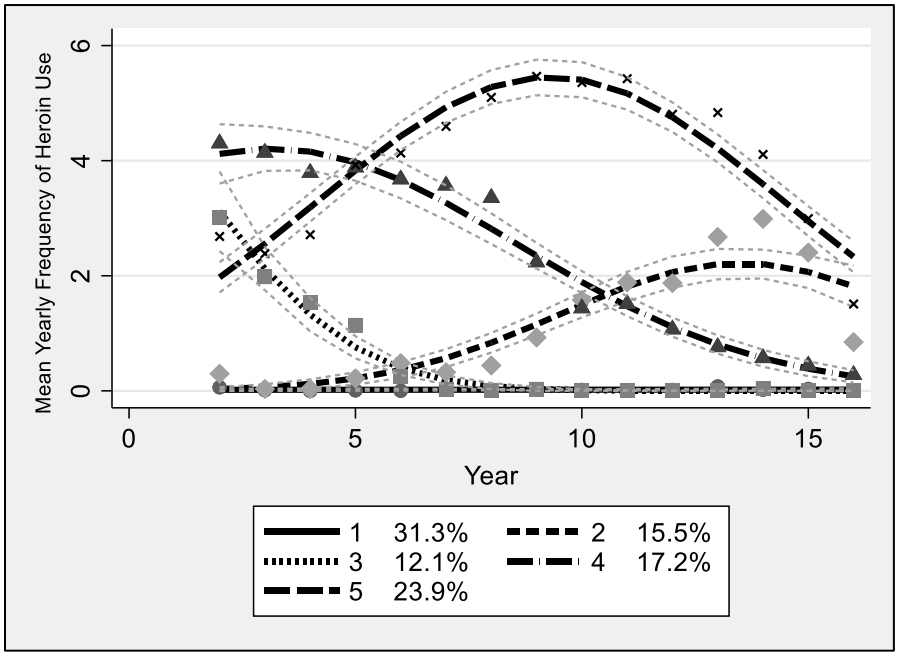

Figure A2. Five-Group Model

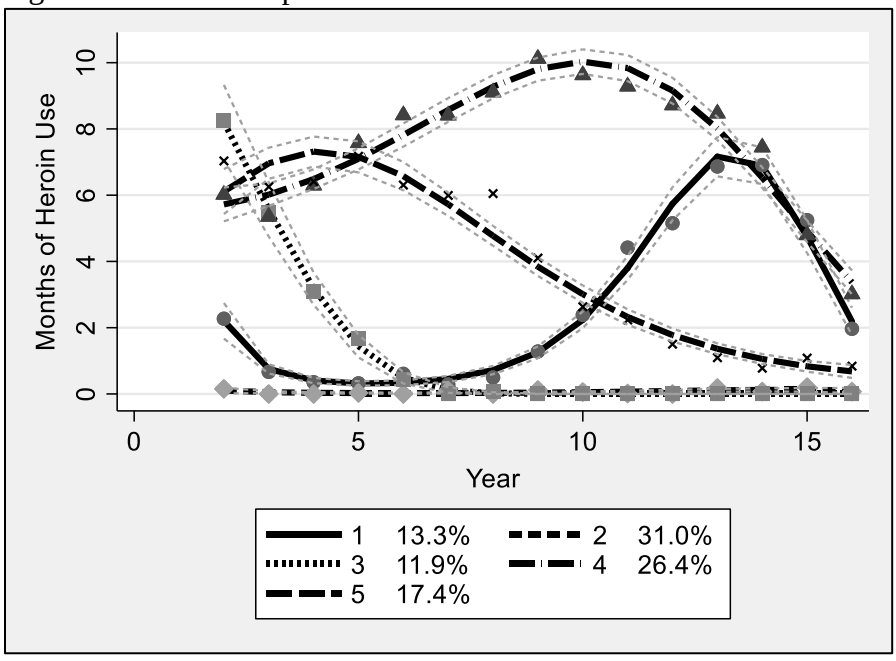

Figure A3. Six-Group Model